**Kozyra J, Ceccarelli A, Torelli E, Lopiccolo A, Gu J, Fellermann H, Stimming U, Krasnogor N.**

**Designing uniquely addressable bio-orthogonal synthetic scaffolds for DNA and RNA origami.**

***ACS Synthetic Biology***

**2017**

**DOI link: https://doi.org/10.1021/acssynbio.6b00271**

ACS Synthetic Biology

Research Article

pubs.acs.org/synthbio

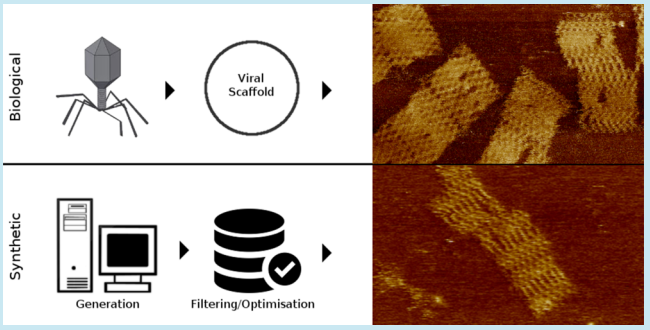
# Designing Uniquely Addressable Bio-orthogonal Synthetic Scaffolds for DNA and RNA Origami

Jerzy Kozyra,[†,‡] Alessandro Ceccarelli,[†,‡,¶] Emanuela Torelli,[†,‡,¶] Annunziata Lopiccolo,[†,‡,¶] Jing-Ying Gu,[§] Harold Fellermann,[*,†,‡] Ulrich Stimming,[§,‡] and Natalio Krasnogor[*,†,‡,¶]

[†]Interdisciplinary Computing and Complex BioSystems (ICOS), School of Computing Science, [‡]Centre for Synthetic Biology and the Bioeconomy (CSBB), [¶]Centre for Bacterial Cell Biology (CBCB), and [§]School of Chemistry, Newcastle University, Newcastle upon Tyne, NE1 7RU, U.K.

*S* *Supporting Information*

**ABSTRACT:** Nanotechnology and synthetic biology are rapidly converging, with DNA origami being one of the leading bridging technologies. DNA origami was shown to work well in a wide array of biotic environments. However, the large majority of extant DNA origami scaffolds utilize bacteriophages or plasmid sequences thus severely limiting its future applicability as a bio-orthogonal nanotechnology platform. In this paper we present the design of biologically inert (i.e., "bio-orthogonal") origami scaffolds. The synthetic scaffolds have the additional advantage of being uniquely addressable (unlike biologically derived ones) and hence are better optimized for high-yield folding. We demonstrate our fully synthetic scaffold design with both DNA and RNA origamis and describe a protocol to produce these bio-orthogonal and uniquely addressable origami scaffolds.




**KEYWORDS:** *DNA origami, sequence design and optimization, nanotechnology, synthetic biology*

Scaffolded DNA origami is a powerful one-pot self-assembly technique[1] that enables construction of custom-shaped objects with nanometre precision. The process involves folding a long single-stranded "scaffold" DNA molecule using multiple short oligonucleotide "staple strands" which bind the scaffold at designated places and hold it in place. A range of nanostructures has been conceived: regular[2,3] and curved solids,[4,5] tubes, and channels,[6,7] as well as controllable DNA nanodevices including box containers,[8] dynamic mechanisms,[9,10] and nanorobots.[11] Natural biocompatibility of DNA and RNA nanodevices makes them an attractive candidate for cellular studies. RNA assemblies have been applied *in vivo* to control hydrogen-production pathways.[12] DNA origamis are stable in lysed cells[13] and can easily interface with biomolecules such as proteins[14−16] and peptides.[17] Treatment with purified nucleases do not compromise the stability of DNA origamis.[18] DNA nanostructures can be delivered into the mammalian cells[19,20] and used as, for example, diffusive molecular cargo[21] and cellular delivery system in Human HEK293.[22] Computation *in vivo* using DNA was shown with nanocontrollers of hemolymph cells in Blaberus discoidalis[23] as well as in a strand exchange mechanism in CHO K1.[24]

Nanotechnology based on DNA and RNA is still in the early developmental stage. However, the potential for nanostructures to assemble and function in a programmable way provides an exciting objective, in particular if such structures could be expressed genetically and manufactured in living cells. Weather DNA origami can be efficiently folded *in vivo* remains an open question, and many limiting challenges must be overcome first. One of the main factors restricting the complexity and applicability of DNA origami is the source of the scaffold which is commonly of viral origin. Existing DNA origami is built from a DNA sequence that is not bio-orthogonal as it contains genetic information; for example, it codes viral proteins and is recognized by various restriction enzymes. These inherent biological features are problematic if one tries to express and fold bio-orthogonal origamis that interfere minimally with a cell's machinery. Little research has focused on addressing this issue, as the phage-based scaffolds became easy to obtain and manipulate.[25] Currently, with the exception of the work of Geary et al.,[26] the sequence design and its optimization is restricted to cyclic permutations of the existing viral scaffolds or modifications of scaffold-staples layouts.[18] On the other hand, while Geary et al.[26] present a synthetic sequence optimized for cotranscription, it requires a different sequence for each nanostructure one may want to assemble.

Furthermore, and of concern not only within a synthetic biology context, the repetition of nucleotide sequences in existing scaffolds and staple strands may cause unspecific

**Special Issue:** IWBDA 2016

**Received:** September 30, 2016
**Published:** April 17, 2017

DOI: 10.1021/acssynbio.6b00271
*ACS Synth. Biol.* XXXX, XXX, XXX−XXX

hybridization.[1] The resulting misfoldings (especially kinetic traps) can disrupt the self-assembly process and lead to structural deformations or malfunction of folded nanodevices. The evidence of potential misfoldings was explored by previous study and prevented by a judicious design of the folding funnel.[27] The problem might be also counteracted by the cooperative nature of the folding[28] and strand-displacement reactions as an error-fixing mechanism. All these effects play a role in the self-assembly of DNA origami, but are currently hard to control in pragmatic manner. On top of that, strand-displacement reactions are known to have slower kinetics compared to hybridization[29] which may be a setback in the folding process.

It seems that part of the problem is caused by the lack of rules for effective sequence design[30] and as such was the main focus of our study. The need for a design of biologically neutral DNA sequences was emphasized recently.[31] However, the current computational methodologies are difficult to apply in the context of DNA origami. Our key motivation was to propose a novel approach to eliminate the ambiguity in the scaffold addressability (i.e., where staples bind) and to ensure the resulting scaffold sequences are biologically neutral (i.e., "bio-orthogonal"). Also, we focused our efforts on providing automated biodesign tools allowing for a rapid scaffold prototyping and analysis (as explained in the following section).

## SYNTHETIC SCAFFOLD DESIGN

In this paper we tackle the above issues by exploiting a property of a family of combinatorial objects called De Bruijn sequences (DBSs). More specifically DBSs of order $k$ have no duplicate subsequences of size $k$ or larger, thus rendering them uniquely addressable by design.

The uniqueness property (i.e., lack of repetitions) makes DBSs attractive candidates for addressable scaffolds: any staple binding a specific region of such scaffold is by design complementary only to that specific region. This in principle should favor specific hybridization over any nonspecific one.

The workflow for constructing synthetic DBS scaffolds for DNA origami is automated with a custom-made software; it is available as a plug-in for a popular open-source tool caDNAno.[32] The overall design pipeline is divided into four steps: (i) construction of De Bruijn graph, (ii) filtering of biological sequences, (iii) construction of alternative De Bruijn sequences and (iv) further optimization. Currently, steps (i-iii) are fully automated, while step (iv) is semiautomated as the optimization procedures differ depending on the criteria chosen for specific application.

**Graph Construction.** In the first step an algorithm constructs a De Bruijn graph that will be later used to produce a synthetic DBS scaffold. For DNA it is possible to construct DBS scaffolds of length $4^k$, where $k$ is the order of the underlying De Bruijn graph.[33] For that reason, one should first determine the length of the scaffold required to conceive the target DNA origami shape. For example, the lexicographically least DBS built with codons (i.e., 3-mers) is 64-nt long (Figure 1). Here, we decided to build a proof-of-concept synthetic scaffold of a size similar to the pUC19 cloning vector. The shortest DBS satisfying this requirement can be built from 6-mers (i.e., DBS of order 6) and thus have a total length of 4096 nt (i.e., $4^6$.)

**Graph Filtering.** A DBS may contain some undesirable sequences, such as restriction enzymes binding sites. This is very likely to occur as our DBS contains all possible 6 nt long

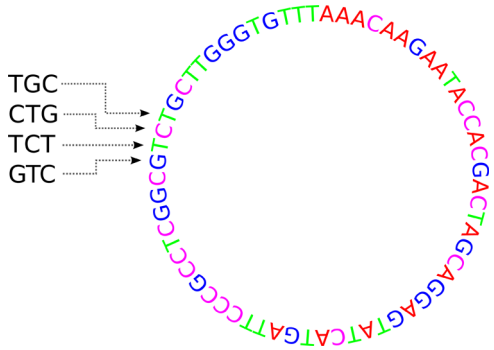



**Figure 1.** Lexicographically least De Bruijn sequence (DBS) of order 3 contains all nucleotide triplets (i.e., codons). De Bruijn sequences are uniquely addressable because repetitions of subsequences are not allowed. Careful observation shows that no three consecutive nucleotides appear more than once, for example, sequences TGC, CTG, TCT, and GTC are unique.

sequences and, as such, it is not bio-orthogonal. One may want to constrain the scaffold construction with certain site-specific sequences. Thus, in the second step, the user specifies a set of forbidden DNA sequences which will be excluded from the DBS scaffold. To demonstrate this, we fetched the sequence data related to *E. coli* K12 from the PRODORIC[34] database together with a list of restriction endonucleases provided by New England Biolabs (NEB). The removal of biologically active sequences should in principle diminish context-dependence and allow the origami system to be isolated *in vivo*.

**Sequence Construction.** Constructing an instance of a DBS is equivalent to finding an Eulerian cycle in a given De Bruijn graph. For an alphabet with four symbols, there are $(4!)^{4^{k-1}} \times 4^{-k}$ distinct cycles, each yielding a distinct sequence. Efficient algorithms for the construction of DBSs exist and were discussed recently in the context of genome assembly.[35] Our software, based on a similar approach, creates DBSs stochastically (i.e., Eularian cycles are picked at random). (For more comprehensive analysis of the underlying graph theory guiding our synthetic scaffold construction and the related software, see Supporting Information.) Here, we generated two DBSs tailored for folding of a square DNA origami (2.4 Kb) and a triangle RNA-DNA hybrid origami (1 Kb).

**Sequence Optimization.** The DBS design space is huge (i.e., practically unlimited for long sequences), affording further optimization. In the final step, the generated sequences can be selected and optimized to a particular specification. Thus, we specified additional criteria to improve the scaffold folding properties. Namely, the DNA origami scaffold (2.4 knt) was picked as a compromise between the following factors: elimination of all forbidden sequences (Table 1) and the stability of secondary structure and nucleotide composition. The ViennaRNA package[36] was used to predict minimum free energy (MFE) of DNA sequences using energy parameters provided by ref 37. The MFE of the pUC19 scaffold is −414.6 kcal/mol (GC content: 0.52). Interestingly, it is more stable than that of a randomly generated sequence with the same nucleotide composition (Figure 2 in red). We aimed to obtain a weaker secondary structure in the DBS scaffold, while preserving similar nucleotide composition (MFE of −376.4 kcal/mol, GC content: 0.5). This should facilitate the origami folding as the scaffold will hybridize more readily with the


DOI: 10.1021/acssynbio.6b00271
ACS Synth. Biol. XXXX, XXX, XXX−XXX

**Table 1. Number of Hits in Databases for Selected Scaffold Sequences**[a]

| scaffold | | NEB | |
|---|---|---|---|
| (length in knt) | PRODORIC | common | all |
| DBS (1) | 0 | 0 | 27 |
| DBS (2.4) | 0 | 0 | 63 |
| pUC19 (2.6) | 28 | 9 | 66 |
| M13mp18 (7.2) | 42 | 9 | 89 |
| λ-phage (48.5) | 69 | 12 | 145 |

[a]The data is from PRODORIC database for *E. coli* strain K12 (1686 entries), NEB list of restriction endonucleases (280 entries; 13 selected as common). Note that both PRODORIC and common NEB were used to constrain the generation of DBSs and thus no hits were found for those databases.

staples rather than with itself. In addition, in RNA−DNA origami both the DBS scaffold (1 knt) as well as the staple sequences were optimized to weaken secondary structures and avoid hairpin formation (see Supporting Information).

To further validate the bio-orthogonality, we used the *Reciprocal Best Hits* (RBH) method. NCBI's BLAST has been used to find alignments of DBSs against known genetic sequences. Significant hits were found when adjusting advanced options of BLASTN to word size 16. The analysis revealed six alignments in the two sequences we designed for this study (Table 2). The low scores confirm the synthetic nature of the DBSs thus further supporting it as a novel bio-orthogonal method for designing DNA origami.

Properties of DBSs provide the DNA origami designer with a rich combinatorial space from where to draw uniquely addressable bio-orthogonal scaffolds. Moreover, the large number of possible DBSs is an essential advantage as it allows for a vast library of bio-orthogonal scaffolds selected for desired biological properties and tailored to particular origami designs. This method provides a necessary alternative to viral sequences and has the potential to find many novel applications at the interface between nanotechnology and synthetic biology.

## LABORATORY VALIDATION

The first synthetic DBS was constructed to fold into a square DNA origami, roughly 50 nm in size, which required 2.4 knt of the scaffold. The shortest DBS satisfying this requirement can be built from subsequences of 6 nt (i.e., DBS of order 6) and thus have a total length of 4096 nt (i.e., $4^6$). However, this theoretical maximum was reduced to 3.3 knt when the DBS was constrained with biological sequences (PRODORIC and NEB) and then trimmed to the length required by the square design.

We first produced a ssDNA scaffold from a double-stranded plasmid as illustrated in Figure 3. The 2.4 Kb De Bruijn DNA sequence encoded in a commercial plasmid was amplified through a PCR. The reverse primer was modified with a biotin molecule linked through a triethylene glycol (TEG) spacer-arm (IDT). The sequence was then attached to magnetic beads and denatured with NaOH. The complementary strand was finally removed via magnetic separation. The single-stranded sequence was purified through agarose gel electrophoresis.[38] The De Bruijn origami was then folded into a square following the rapid isothermal protocol described by Sobczak et al.[28] This method grants a more stable product with a lower rate of misfolding, reducing the folding time from hours to minutes. The samples were finally analyzed by AFM to compare the quality against the pUC19 DNA origami (Figure 4).

As a control experiment, we constructed a similar origami shape using the pUC19 scaffold which was prepared according to the previous study.[2] In short, a nicking enzyme followed by exonuclease treatment removed the "antiscaffold" strand and left the scaffold intact which was then folded into the square DNA origami shape (see Methods). Because the folding of a square required only 2.4 knt of scaffold, the remaining 200 nt were left unpaired and formed a dangling loop at the corner of the shape (Figure 5).

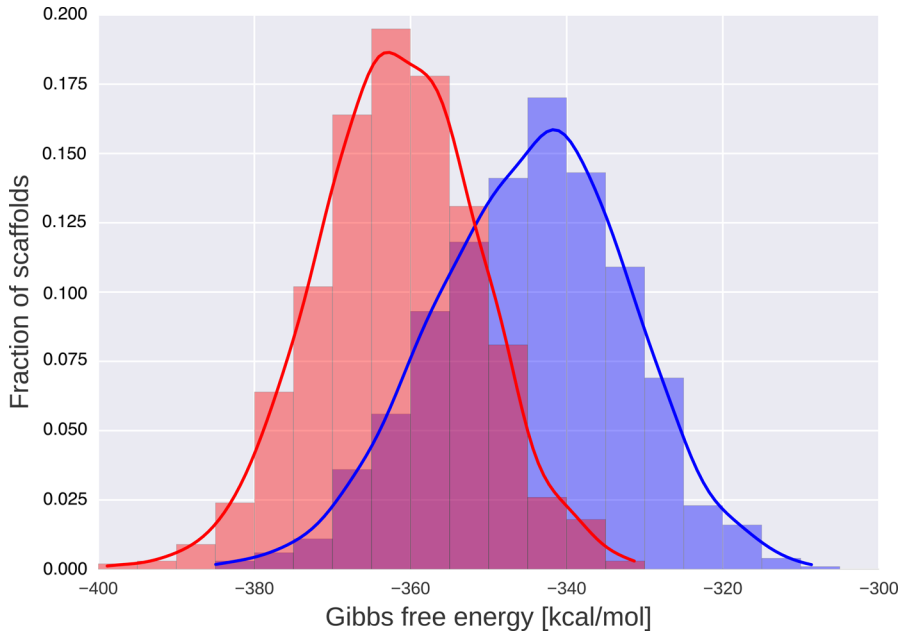



**Figure 2.** MFE distribution: 1000 DBS scaffolds (blue) and 1000 randomly generated scaffolds with similar nucleotide composition as pUC19 (red), are plotted according to the secondary structure energy. Sequences are 2.4 knt long.



DOI: 10.1021/acssynbio.6b00271
*ACS Synth. Biol.* XXXX, XXX, XXX−XXX

**Table 2. BLAST Alignment Results**

| scaffold | genome | score | cover | E value | identities | accession no. |
|---|---|---|---|---|---|---|
| DBS (1) | Halichoerus grypus | 43.6 | 2% | 7.6 | 100% | JX218922.1 |
| DBS (2.4) | Spirometra erinaceieuropaei | 51.0 | 1% | 0.11 | 100% | LN056044.1 |
| DBS (2.4) | Thelazia callipaeda | 51.0 | 1% | 0.15 | 86% | LK979655.1 |
| DBS (2.4) | Ovis canadensis canadensis | 47.3 | 1% | 1.4 | 94% | CP011893.1 |
| DBS (2.4) | Ovis Aries (predicted) | 47.3 | 1% | 1.4 | 94% | XR_001042372.1 |
| DBS (2.4) | Aerococcus sanguinicola | 47.3 | 1% | 2.0 | 100% | CP014160.1 |
| DBS (2.4) | Protopolystoma xenopodis | 45.4 | 1% | 5.2 | 96% | LM741928.1 |
| DBS (2.4) | Macaca fascicularis | 45.4 | 1% | 7.2 | 89% | LT160001.1 |

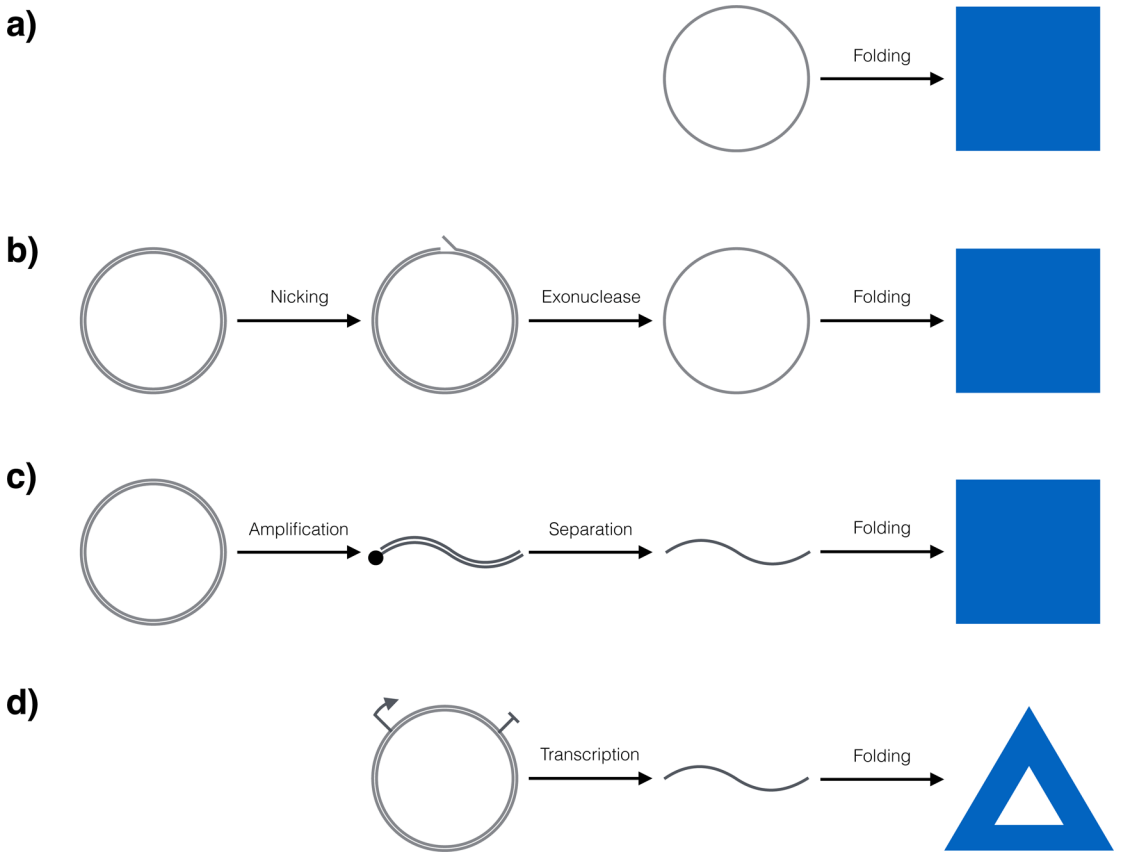



**Figure 3.** Scaffold preparation protocols: (a) the common folding protocol assumes ready-to-use viral ssDNA; (b) removal of antiscaffold strand from pUC19 via enzymatic reactions; (c) removal of antiscaffold strand from De Bruijn PCR product using magnetic beads; (d) transcription of RNA De Bruijn scaffold.

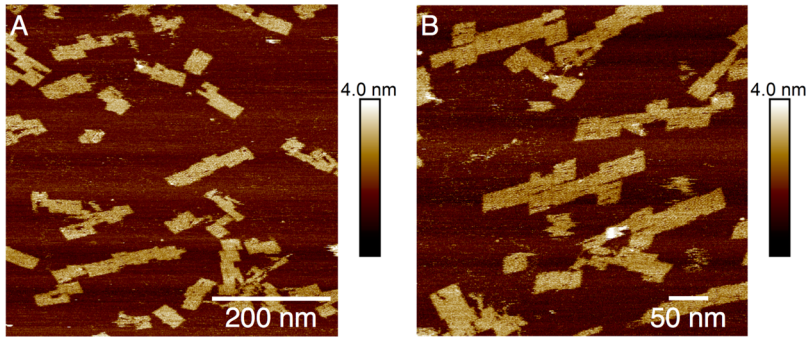



**Figure 4.** AFM imaging of the square DNA origami based on a DBS scaffold.

Furthermore, to test our design for RNA−DNA hybrid origamis, we constructed a 1 knt long DBS scaffold which was designed to fold into a triangle with a hole. Again, a DBS satisfying this length requirement was built from 6 nt long unique subsequences. The DBS was additionally constrained to exclude RNA-specific sequences: the start codon, Shine−



DOI: 10.1021/acssynbio.6b00271
ACS Synth. Biol. XXXX, XXX, XXX−XXX

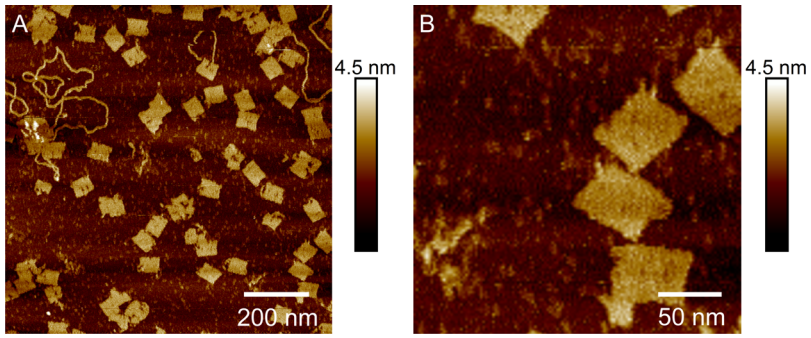



**Figure 5.** AFM imaging of the square DNA origami based on the pUC19 scaffold.

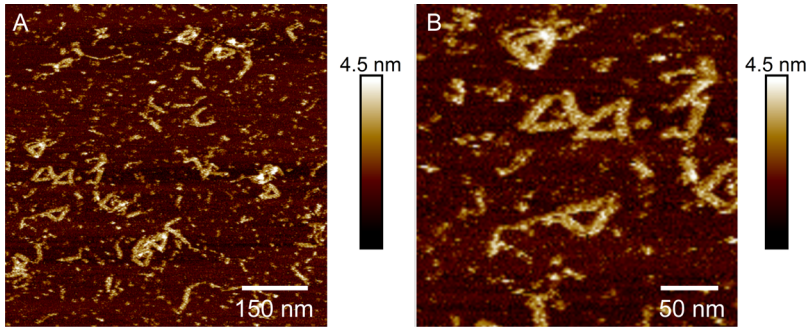



**Figure 6.** AFM imaging of the triangle RNA–DNA hybrid origami based on a DBS scaffold.

Dalgarno sequence, and four additional restriction enzymes. The T7 promoter was inserted upstream of the DBS (Figure 3d) to enable transcription *in vitro* similar to the previous study.[39] The folding was performed following previous studies[39] and AFM imaging confirmed the correct folding of the triangle structures (Figure 6). These two experiments demonstrate that DBS scaffolds can be utilized in the same manner as viral ones without change of the folding protocol.

An additional experiment was carried out to demonstrate the bio-orthogonality of the scaffold based on fluorophore labeling with electroporation. The DBS (1 Kb) scaffold was successfully labeled using the Ulysis Nucleic Acid Labeling kit (Figure 7a). After purification, the labeled scaffold was added to *E. coli* cells via electroporation without preincubation for 10 min. When the preincubation influence was tested, a lower band intensity was recorded (Figure 7b, lanes 5 and 8). For this reason, the subsequent experiments were performed without preincubation (Figure 7c). *E. coli* cell suspension was recovered for 25 or 40 min, embedded in agarose plugs, lysed, and loaded into the well of agarose gels. To assess the resistance of electroporated scaffold inside the cells, samples were analyzed using agarose gel electrophoresis and visualized under a Typhoon laser scanner. Figure 7 panels b and c show that the Alexa 488 labeled scaffold was not degraded inside the cells.

## FURTHER CHARACTERIZATION

We analyzed statistical redundancy of the three common scaffolds for DNA origami: pUC19, M13mp18, and $\lambda$-phage. The number of the repeated sequences ($k$-mers) was determined using a suffix tree analysis and plotted according to the length $k$ (Figure 8a). Note that only the longest repeats are shown and their respective subsequences are excluded. Existing scaffolds contain many repetitions which are longer than the typical binding domains of staples. In 2-dimensional structures (Figure 8b), staple domains are usually composed of 8 nt (or multiples of it). In 3-dimensional structures (build on honeycomb lattice) domains are shorter—typically multiples of 7 nt. For example, the most frequently used scaffold, M13mp18 has over $10^3$ repeats of length $\geq$8 nt, while $\lambda$-phage has over $10^4$ of them. How many of these repeats occur at staple binding domains depends on the particular design and choice of the corresponding staple set. Generally, the number as well as length of repeating sequences grows proportionally to the scaffold length. M13mp18, has the longest repeats spanning 29, 30, and 42 nt which are representative examples of ambiguity in staple addressability. Interestingly, they appear as outliers in the underlying distribution of repeats and are not present in the other two scaffolds, for which the longest repeats span 13 and 15 nt, respectively. In comparison synthetic DBS scaffolds of length 4 kb, 16 kb, and 64 kb can be constructed such that they have no repeats longer than 5 nt, 6 nt, and 7 nt, respectively.

Sequence uniqueness alone does not guarantee binding at a unique target location because a scaffold might contain slight alternations of the sequence to which a staple might bind (although with smaller binding affinity due to mismatches). We therefore investigated unique addressability based on binding energies in different designs and scaffolds configurations. A custom-built algorithm was used to calculate the addressability measure for each staple. First a simple heuristic based on Levenshtein distance finds all possible regions of the scaffold to which a staple can hybridize. When a possible binding site is detected the associated thermodynamic potential is derived using the ViennaRNA package[36] (with appropriate energy parameters provided by ref 37). The resulting thermodynamic potentials (measured as Gibbs free energy) are used to establish the relative probability of a staple hybridizing at a specific location according to the Boltzmann distribution. In other words, as the addressability measure decreases the staple is more likely to bind at an incorrect target.



DOI: 10.1021/acssynbio.6b00271
*ACS Synth. Biol.* XXXX, XXX, XXX–XXX

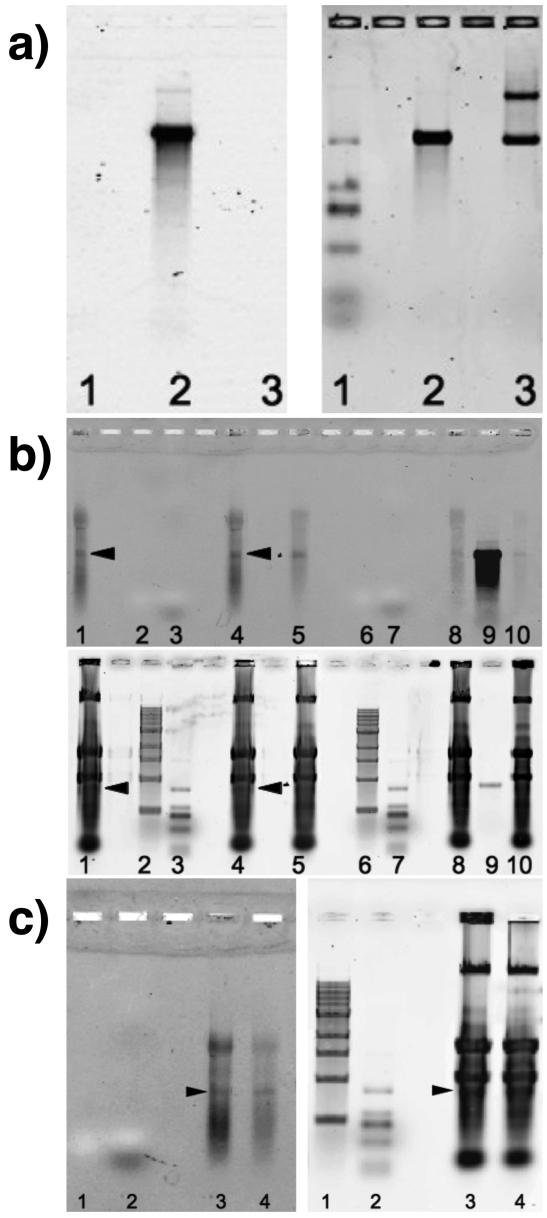



**Figure 7.** Stability of DBS in living cells: (a) Agarose gel electrophoresis of Alexa 488 labeled DBS scaffold imaged before (left) and after (right) ethidium bromide staining. Lane 1, low range ssRNA ladder; lane 2, Alexa 488 labeled scaffold DBS; lane 3, not labeled scaffold DBS. (b) Agarose gel electrophoresis of *E. coli* lysed cells after electroporation with Alexa 488 labeled DBS, before (top) and after (bottom) ethidium bromide staining. Lanes 1 and 4, electroporated *E. coli* with Alexa 488 labeled RNA scaffold (without preincubation); lanes 2 and 6, 1 Kb DNA ladder; lanes 3 and 7, low range ssRNA ladder; lanes 5 and 8, electroporated *E. coli* with Alexa 488 labeled RNA scaffold (preincubation: 10 min); lane 9, Alexa 488 labeled RNA scaffold; lane 10, negative control, not electroporated *E. coli* with Alexa 488 labeled RNA scaffold. (c) Agarose gel electrophoresis of *E. coli* lysed cells after electroporation with Alexa 488 labeled DBS, before (left) and after (right) ethidium bromide staining. Lane 1, 1 Kb DNA ladder; lane 2, low range ssRNA ladder; lanes 3 and 4, electroporated *E. coli* with Alexa 488 labeled RNA scaffold after 40 min of recovery. Black arrows show Alexa 488 labeled scaffold.

Interestingly, we found that longer staple domains (>8 nt long) have nearly perfect addressability measure (the probability is approximately 1) in all examined designs. For short domains (≤8 nt long) there is a strong tendency for DBS scaffolds to have a higher addressability measure than their biological counterparts (Figure 8c). It is the case not only for pUC19 and DBS (2.4 knt) which fold into a small DNA origami tile (presented in this study), but also for the theoretical medium tile design (85 × 85 nm) based on M13mp18 and DBS (order 7). Although the addressability measure in the large tile designs (200 × 200 nm) is generally low, the synthetic DBS (order 8) still outperforms the λ-phage scaffold. These results suggest that longer scaffolds have a higher probability of mismatching, which is partly caused by repeats in scaffold, and therefore ambiguity of staple addressing. Moreover, it might explain the difficulties of folding larger DNA origami using the λ-phage scaffold.

## DISCUSSION

We propose synthetic scaffolds which widen the programmability of DNA origamis and ensure their bio-orthogonality. We show that synthetic DBS scaffolds which are uniquely addressable and bio-orthogonal *by design* fold into DNA origami and RNA−DNA hybrid origami without alteration of the folding protocol.

Our computational analysis shows that the repetition of sequences in natural scaffolds has a negative impact on staple specificity. This problem is magnified for longer scaffolds because the number of potentially stable targets for a staple grows proportionally with the scaffold length. Although an obvious solution might appear to be the use of only long staples, the exact hybridization kinetics of longer sequences is poorly understood while sparse double-crossover motifs may compromise the rigidity of nanostructures. Thus, the use of natural sequences does not scale well for creation of large objects based on a single scaffold. Further, we show that scaffolds based on DBS provide more specificity and are therefore uniquely addressable.

Establishing the nonspecific sources of interactions (or interference) within a given biological system is a challenging task. Regardless of this difficulty, we explained here how our method allows explicit exclusion or inclusion of sequence specific sites from the final DNA nanostructure. This new method allows to strictly control the interface between the origami devices and various biomolecules through the insertions of biological sites in otherwise bio-orthogonal scaffolds. Generalisation of the origami programmability may substantially reduce the barrier to entry of this methodology to biological applications with the posibility of extreme-precision patterning as well as self-assembly in biotic environments.

Future efforts should concentrate on two goals: experimental confirmation of the superior folding for long scaffolds as well as development of protocols allowing folding *in vivo*.

## METHODS

**Software.** The computer code for generating uniquely addressable, bio-orthogonal synthetic scaffold sequences is available online (see Supporting Information).

**DNA Origami and RNA−DNA Hybrid Origami Designs.** The DNA origami square and the RNA−DNA hybrid origami triangle (see Supporting Information Figures) have been designed using caDNAno as described elsewhere.[1,39]

**DNA Origami Folding.** Briefly, the pUC19 plasmid was treated with Nt.BspQI nicking enzyme (NEB, UK) at 50 °C for


DOI: 10.1021/acssynbio.6b00271
ACS Synth. Biol. XXXX, XXX, XXX−XXX

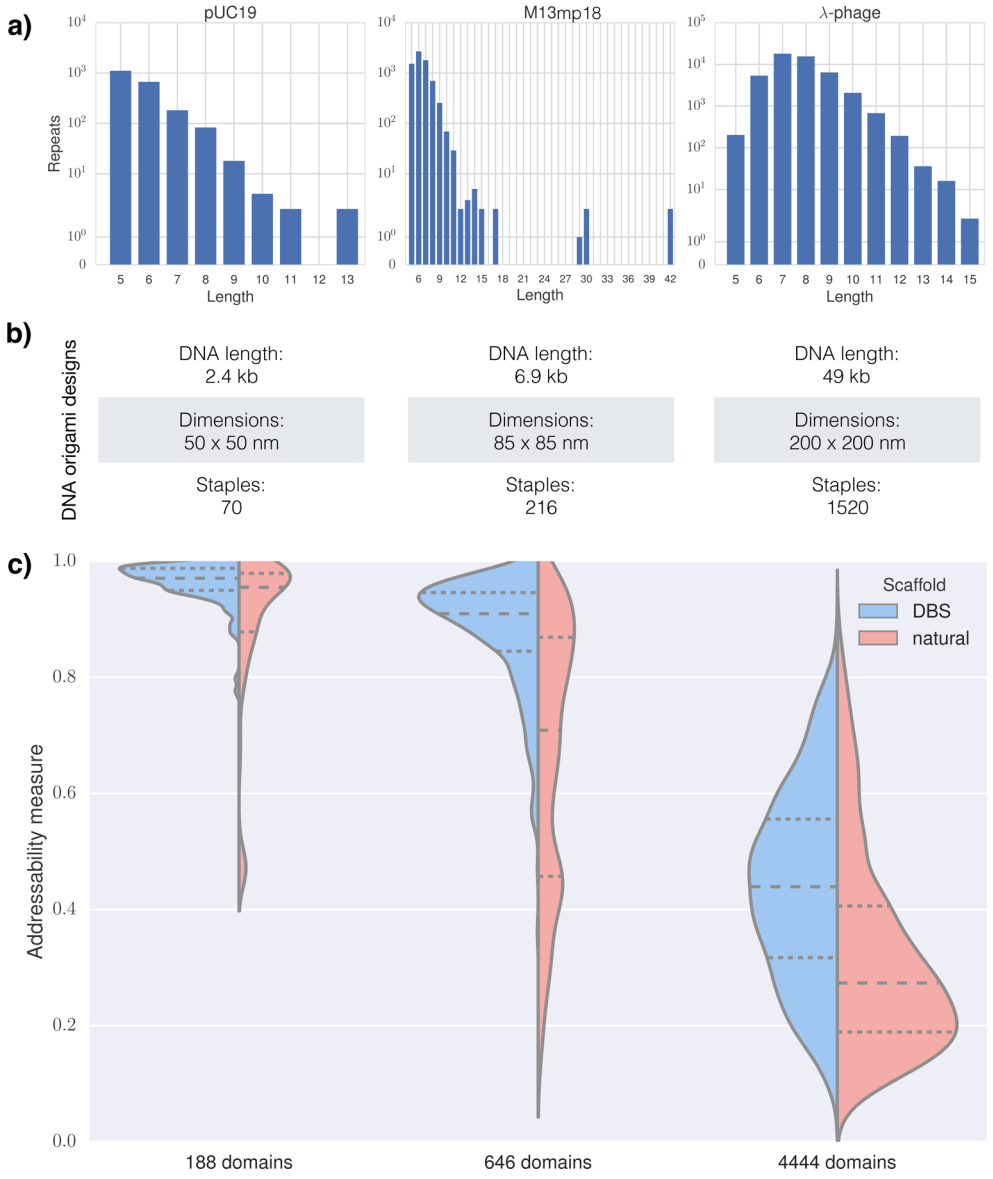



**Figure 8.** Characterization of DNA origami scaling. (a) Sequence repetitions in three scaffolds: 2.6 kb pUC19 vector (left), 7.2 kb M13mp18 bacteriophage genome (middle), and 48 kb λ-phage genome (right). (b) Comparison of simple 2-dimensional square tile designs conceived with respective scaffolds. (c) Thermodynamic addressability in respective designs: the violin plots show the probability distribution of the staples hybridizing with specific (i.e., designed) domains in the scaffold. The dashed lines mark the quartiles of each distribution. For clarity only the short staple domains (8 nt in length) are shown, as the probability measure for longer domains is approximately 1.

90 min. The mix solution was then incubated for 20 h at 37 °C with T7 exonuclease and Lambda exonuclease to remove the complementary strand, leaving the scaffold intact (Figure 3b). The ssDNA scaffold was ethanol precipitated, air-dried and then dissolved in Tris-EDTA (TE) buffer.

To generate a linear pUC19 scaffold, a short oligo (GCCACCTGACGTCTAAGAAA) which contains a restriction enzyme site (underlined), ZraI, was designed and synthesized. The circular single-strand pUC19 was then incubated with this oligo and treated with ZraI at 37 °C for 45 min. After heat inactivation, the digested DNA was then purified and concentrated by ethanol precipitation and resuspended in TE buffer as linear single-strand scaffold.

The 2.4 Kb DBS was synthesized and cloned into a plasmid commercially (Life Technologies, UK). To generate linear ssDNA, a PCR based method with 3′ biotinylated reverse primer was used as in previous study[38] with necessary modifications. The biotinlyated PCR product was captured by streptavidin coated magnetic beads (Dynabeads kilobase-BINDER kit). After the treatment with 0.2 M NaOH, the (ssDNA) scaffold strand was released and subsequently neutralized by 3 M $NH_4(OAc)$. The product was then purified using a PCR purification kit (Qiagen) to remove the salts and agarose gel electrophoresis (1%) to separate the scaffold from the residual double stranded DNA. The ssDNA was recovered using a Freeze'n'squeeze gel extraction kit (BioRAD) and finally purified with a PCR purification kit. For the assembly reaction, 10 nM ssDNA scaffold and 200 nM each staples oligos were mixed in a folding buffer containing 5 mM Tris, 5 mM NaCl, 1 mM EDTA (pH of 8) and 8 mM $Mg(OAc)_2$ (pH of 8). The reaction was heated to 95 °C for 30 s and cooled to 51 °C for 10 min in a thermal cycler, following.[28] The staple excess was removed using Amicon Ultra centrifugal filters 100 kDa.



DOI: 10.1021/acssynbio.6b00271
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Electrophoresis of the folded DNA was carried out in 2% agarose gel containing 0.5 μg/mL ethidium bromide and 1× TBE/Mg buffer (40 mM Tris, 4 mM Acetate, 1 mM EDTA). The electrophoresis gels were run for 2 h at 70 V in an ice/water-cooled tray. The DNA bands in gels were visualized using ultraviolet light and the desired band was excised by scalpels. The DNA in excised gels was then extracted using Bio-Rad freeze'n'squeeze column according to manufacturer's instruction. The excess of staples was removed using Amicon Ultra 100 kDa. The recovered material was then prepared for imaging.

**RNA−DNA Hybrid Origami Folding.** The 1 Kb DBS preceded by a T7 promoter was synthesized and cloned in a commercial 14AA575P plasmid (Life Technologies, UK). The DNA template-scaffold was obtained by PCR amplification with Phusion Hi-fidelity DNA Polymerase (NEB, UK). The RNA scaffold was synthesized using Ampliscribe T7-Flash Transcription kit (Epicenter) on the DNA scaffold template at 42 °C for 100 min. The scaffold was subsequently purified through a phenol-chloroform-isoamyl (125:24:1 Sigma-Aldrich) and chloroform (Sigma-Aldrich) precipitation. The concentration of the nucleic acids was evaluated by Nanodrop analysis (Thermo scientific).

Folding of the origami was carried out in TAE buffer (40 mM Tris, 4 mM Acetate, 1 mM EDTA) enriched with magnesium acetate 12.5 mM. The reaction was performed using a concentration of 10 nM of RNA scaffold and 100 nM of DNA staples oligos. The folding solution was incubated for 10 min at 65 °C followed by a temperature ramp of 0.01 °C/s to 25 °C and maintained at that temperature for 5 min. The solution was then held at constant temperature of 4 °C to stop the reaction. The origami structures were purified through Amicon ultra filters 100 kDa to remove the excess of free staples and to concentrate the samples.

**Atomic Force Microscope (AFM).** Scanning Probe Microscopies (SPM) can have a high ability to image organic structures such as proteins under in situ conditions as has been shown for scanning tunnelling microscopy (STM) and scanning electrochemical potential microscopy (SECPM).[40] Here, AFM has been used which has a lower resolution but reflects more accurately the morphology of the imaged object.

For AFM imaging, 10 μL of nickel acetate tetrahydrate 10 mM was applied onto freshly cleaved mica to stabilize the sample on the substrate. Subsequently, 5 μL of purified origami sample solution was applied and let to airdry. The same origami folding buffer was used as imaging buffer. AFM imaging was performed on a Bruker multimode 8 AFM in Scanasyst mode, using Bruker Scanasyst-Fluid+ tip.

**Alexa 488 Labeled DBS Scaffold Preparation.** Labeling of the DBS scaffold was performed using Ulysis Nucleic Acid Labeling kit (Life Technologies) following manufacturer's protocol. Briefly, after denaturation at 95 °C for 5 min of the purified RNA transcript (1 μg), the Alexa Fluor 488 ULS labeling reagent stock solution (1 μL) and the labeling buffer were added to the tube containing the denatured RNA sample (final volume 25 μL). The reaction was incubated at 90 °C for 10 min and stopped by plunging the tube into an ice bath. Two labeled RNA solutions (50 μL) were purified by using Micro Bio-Spin P-30 spin column (BioRad) and eluted in 10 mM Tris-HCl pH 7.4.

To check the labeling reaction, the Alexa 488 labeled scaffold was run on 2% agarose gel at 100 V for 1 h. After the electrophoresis was completed, the gel was imaged using Typhoon laser scanner (excitation 488 nm, emission 532 nm; GE Healthcare Life Sciences). Then the gel was then stained with an ethidium bromide solution and the band was observed using a Typhoon laser scanner.

**Electroporation and Plug Preparation.** The commercial electrocompetent bacterial cell line used for electroporation was NEB 5-alpha electrocompetent *E. coli* (NEB). Cells were diluted 1:1 with sterile milli-Q water in a prechilled tube and 20 μL of cells suspension were used for each electroporation experiment after the addition of 5 μL of labeled scaffold DBS. The mixture of competent cells and labeled molecules was immediately transferred into a prechilled electroporation cuvette (0.2 cm gap cuvette, BioRad) and placed into an electroporator (Gene Pulser, BioRad). The electroporation conditions were 1.8 kV, 200 Ω, 25 μF. The negative control was prepared incubating cells with the same volume of the same labeled scaffold solution: the cells were not electroporated and were washed as the electroporated samples. Immediately 500 μL of 37 °C SOC medium was added to the cuvette, gently mixed up and down twice and transferred to culture tubes. Cells were allowed to recover for 25 or 40 min at 37 °C under shaking. After recovery, cells were harvested by centrifugation at 3300*g* for 1 min at 4 °C, washed 5 times with 500 μL of 1× phosphate buffered saline solution pH 7.4 (PBS, Chem Cruz) and resuspended in 50 μL of PBS.[41,42] A 2.5 μL portion of proteinase K (20 mg/mL stock, NEB) and 50 μL of melted SeaKem Gold agarose (Lonza) in TE (10 mM Tris, 1 mM EDTA pH 8.0) were added to the cell suspension and mixed gently. The agarose-cell suspension mixture (102.5 μL) was immediately dispensed into a well of plug mold (BioRad). Each agarose plug was allowed to solidify at 4 °C for 20 min.[43]

**Lysis of Cells in Plugs.** The plugs were incubated in 5 mL of cell lysis buffer (50 mM Tris, 50 mM EDTA pH 8, 1% sarcosine, 0.1 mg of proteinase k/mL) for 15 min at 54 °C in a water bath.[43] After lysis, the plugs were washed four times (10 min/wash) at room temperature (once with nucleases free water and three times with TE pH 8.0). After washing steps the plug slices were loaded into the wells of 1% agarose gel: the electrophoresis was performed for 1 h at 100 V. The 1 Kb DNA ladder (NEB) and the low range ssRNA ladder (NEB) were used as molecular weight markers. The gels were imaged using Typhoon laser scanner (excitation 488 nm, emission 532 nm). After staining with ethidium bromide solution, the gels were scanned again on Typhoon laser scanner.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.6b00271.

DNA origami and RNA−DNA origami designs; software; sequence generation algorithms; scaffold and staple sequences (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

*E-mail: harold.fellermann@newcastle.ac.uk.
*E-mail: natalio.krasnogor@newcastle.ac.uk.

### ORCID

Harold Fellermann: 0000-0001-5861-1945
Natalio Krasnogor: 0000-0002-2651-4320


### Notes

The authors declare no competing financial interest.



DOI: 10.1021/acssynbio.6b00271
ACS Synth. Biol. XXXX, XXX, XXX−XXX

## ■ ACKNOWLEDGMENTS

This work was supported by EPSRC Grant Nos. EP/J004111/2, EP/L001489/2, and EP/N031962/1.

## ■ REFERENCES

(1) Rothemund, P. W. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature 440*, 297−302.
(2) Douglas, S. M., Dietz, H., Liedl, T., Hogberg, B., Graf, F., and Shih, W. M. (2009) Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature 459*, 414−418.
(3) Martin, T. G., and Dietz, H. (2012) Magnesium-free self-assembly of multi-layer DNA objects. *Nat. Commun. 3*, 1103.
(4) Dietz, H., Douglas, S. M., and Shih, W. M. (2009) Folding DNA into Twisted and Curved Nanoscale Shapes. *Science 325*, 725−730.
(5) Han, D., Pal, S., Nangreave, J., Deng, Z., Liu, Y., and Yan, H. (2011) DNA Origami with Complex Curvatures in Three-Dimensional Space. *Science 332*, 342−346.
(6) Yin, P., Hariadi, R. F., Sahu, S., Choi, H. M. T., Park, S. H., LaBean, T. H., and Reif, J. H. (2008) Programming DNA Tube Circumferences. *Science 321*, 824−826.
(7) Langecker, M., Arnaut, V., Martin, T. G., List, J., Renner, S., Mayer, M., Dietz, H., and Simmel, F. C. (2012) Synthetic Lipid Membrane Channels Formed by Designed DNA Nanostructures. *Science 338*, 932−936.
(8) Andersen, E. S., Dong, M., Nielsen, M. M., Jahn, K., Subramani, R., Mamdouh, W., Golas, M. M., Sander, B., Stark, H., Oliveira, C. L. P., Pedersen, J. S., Birkedal, V., Besenbacher, F., Gothelf, K. V., and Kjems, J. (2009) Self-assembly of a nanoscale DNA box with a controllable lid. *Nature 459*, 73−76.
(9) Marras, A. E., Zhou, L., Su, H.-J., and Castro, C. E. (2015) Programmable motion of DNA origami mechanisms. *Proc. Natl. Acad. Sci. U. S. A. 112*, 713−718.
(10) Fellermann, H., Lopiccolo, A., Kozyra, J., and Krasnogor, N. (2016) In Vitro Implementation of a Stack Data Structure Based on DNA Strand Displacement. *Unconventional Computation and Natural Computation 9726*, 87−98.
(11) Douglas, S. M., Bachelet, I., and Church, G. M. (2012) A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science 335*, 831−834.
(12) Delebecque, C. J., Lindner, A. B., Silver, P. A., and Aldaye, F. A. (2011) Organization of Intracellular Reactions with Rationally Designed RNA Assemblies. *Science 333*, 470−474.
(13) Mei, Q., Wei, X., Su, F., Liu, Y., Youngbull, C., Johnson, R., Lindsay, S., Yan, H., and Meldrum, D. (2011) Stability of DNA Origami Nanoarrays in Cell Lysate. *Nano Lett. 11*, 1477−1482.
(14) Saccà, B., Meyer, R., Erkelenz, M., Kiko, K., Arndt, A., Schroeder, H., Rabe, K. S., and Niemeyer, C. M. (2010) Orthogonal Protein Decoration of DNA Origami. *Angew. Chem., Int. Ed. 49*, 9378−9383.
(15) Park, S. H., Yin, P., Liu, Y., Reif, J. H., LaBean, T. H., and Yan, H. (2005) Programmable DNA Self-Assemblies for Nanoscale Organization of Ligands and Proteins. *Nano Lett. 5*, 729−733. PMID: 15826117.
(16) Lund, K., Liu, Y., Lindsay, S., and Yan, H. (2005) Self-Assembling a Molecular Pegboard. *J. Am. Chem. Soc. 127*, 17606−17607. PMID: 16351081.
(17) Williams, B., Lund, K., Liu, Y., Yan, H., and Chaput, J. (2007) Self-Assembled Peptide Nanoarrays: An Approach to Studying Protein-Protein Interactions. *Angew. Chem., Int. Ed. 46*, 3051−3054.
(18) Castro, C. E., Kilchherr, F., Kim, D.-N., Shiao, E. L., Wauer, T., Wortmann, P., Bathe, M., and Dietz, H. (2011) A primer to scaffolded DNA origami. *Nat. Methods 8*, 221−229.
(19) Jiang, Q., Song, C., Nangreave, J., Liu, X., Lin, L., Qiu, D., Wang, Z.-G., Zou, G., Liang, X., Yan, H., and Ding, B. (2012) DNA Origami as a Carrier for Circumvention of Drug Resistance. *J. Am. Chem. Soc. 134*, 13396−13403. PMID: 22803823.
(20) Schüller, V. J., Heidegger, S., Sandholzer, N., Nickels, P. C., Suhartha, N. A., Endres, S., Bourquin, C., and Liedl, T. (2011) Cellular Immunostimulation by CpG-Sequence-Coated DNA Origami Structures. *ACS Nano 5*, 9696−9702. PMID: 22092186.
(21) Kopperger, E., Pirzer, T., and Simmel, F. C. (2015) Diffusive Transport of Molecular Cargo Tethered to a DNA Origami Platform. *Nano Lett. 15*, 2693−2699. PMID: 25739805.
(22) Mikkilä, J., Eskelinen, A.-P., Niemelä, E. H., Linko, V., Frilander, M. J., Törmä, P., and Kostiainen, M. A. (2014) Virus-Encapsulated DNA Origami Nanostructures for Cellular Delivery. *Nano Lett. 14*, 2196−2200. PMID: 24627955.
(23) Amir, Y., Ben-Ishay, E., Levner, D., Ittah, S., Abu-Horowitz, A., and Bachelet, I. (2014) Universal computing by DNA origami robots in a living animal. *Nat. Nanotechnol. 9*, 353−357.
(24) Groves, B., Chen, Y.-J., Zurla, C., Pochekailov, S., Kirschman, J. L., Santangelo, P. J., and Seelig, G. (2016) Computing in mammalian cells with nucleic acid strand exchange. *Nat. Nanotechnol. 11*, 287−294.
(25) Kick, B., Praetorius, F., Dietz, H., and Weuster-Botz, D. (2015) Efficient Production of Single-Stranded Phage DNA as Scaffolds for DNA Origami. *Nano Lett. 15*, 4672−4676. PMID: 26028443.
(26) Geary, C., Rothemund, P. W. K., and Andersen, E. S. (2014) A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science 345*, 799−804.
(27) Dunn, K. E., Dannenberg, F., Ouldridge, T. E., Kwiatkowska, M., Turberfield, A. J., and Bath, J. (2015) Guiding the folding pathway of DNA origami. *Nature 525*, 82−86.
(28) Sobczak, J.-P. J., Martin, T. G., Gerling, T., and Dietz, H. (2012) Rapid Folding of DNA into Nanoscale Shapes at Constant Temperature. *Science 338*, 1458−1461.
(29) Srinivas, N., Ouldridge, T. E., Sulc, P., Schaeffer, J. M., Yurke, B., Louis, A. A., Doye, J. P. K., and Winfree, E. (2013) On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res. 41*, 10641−10658.
(30) Pinheiro, A. V., Han, D., Shih, W. M., and Yan, H. (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol. 6*, 763−772.
(31) Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T., and MacDonald, J. T. (2014) R2oDNA Designer: Computational Design of Biologically Neutral Synthetic DNA Sequences. *ACS Synth. Biol. 3*, 525−528. PMID: 24933158.
(32) Douglas, S. M., Marblestone, A. H., Teerapittayanon, S., Vazquez, A., Church, G. M., and Shih, W. M. (2009) Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res. 37*, 5001−5006.
(33) de Bruijn, N. (1946) A combinatorial problem. *Proc. Nederl. Akad. Wetensch. 49*, 758−764.
(34) Münch, R., Hiller, K., Barg, H., Heldt, D., Linz, S., Wingender, E., and Jahn, D. (2003) PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res. 31*, 266−269.
(35) Compeau, P. E. C., Pevzner, P. A., and Tesler, G. (2011) How to apply de Bruijn graphs to genome assembly. *Nat. Biotechnol. 29*, 987−991.
(36) Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., and Hofacker, I. L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol. 6*, 26−26.
(37) Turner, D. H., and Mathews, D. H. (2010) NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res. 38*, D280−D282.
(38) Pound, E., Ashton, J. R., Becerril, H. A., and Woolley, A. T. (2009) Polymerase Chain Reaction Based Scaffold Preparation for the Production of Thin, Branched DNA Origami Nanostructures of Arbitrary Sizes. *Nano Lett. 9*, 4302−4305. PMID: 19995086.
(39) Wang, P., Hyeon Ko, S., Tian, C., Hao, C., and Mao, C. (2013) RNA-DNA hybrid origami: folding of a long RNA single strand into complex nanostructures using short DNA helper strands. *Chem. Commun. 49*, 5462−5464.
(40) Baier, C., and Stimming, U. (2009) Imaging Single Enzyme Molecules under In Situ Conditions. *Angew. Chem., Int. Ed. 48*, 5542−5544.
(41) Crawford, R., Torella, J. P., Aigrain, L., Plochowietz, A., Gryte, K., Uphoff, S., and Kapanidis, A. N. (2013) Long-Lived Intracellular


DOI: 10.1021/acssynbio.6b00271
*ACS Synth. Biol.* XXXX, XXX, XXX−XXX

Single-Molecule Fluorescence Using Electroporated Molecules. *Biophys. J. 105*, 2439−2450.
(42) Aigrain, L., Sustarsic, M., Crawford, R., Plochowietz, A., and Kapanidis, A. N. (2015) Internalization and Observation of Fluorescent Biomolecules in Living Microorganisms via Electroporation. *J. Visualized Exp.*, e52208.
(43) Ribot, E. M., Fitzgerald, C., Kubota, K., Swaminathan, B., and Barrett, T. J. (2001) Rapid Pulsed-Field Gel Electrophoresis Protocol for Subtyping of *Campylobacter jejuni*. *Journal of Clinical Microbiology 39*, 1889−1894.



DOI: 10.1021/acssynbio.6b00271
*ACS Synth. Biol.* XXXX, XXX, XXX−XXX